Emanuele Rossi

Barcelona emanuele.rossi1909@gmail.com emanuelerossi.co.uk emanuele-rossi emalgorithm

Summary

I'm an ML researcher specializing in generative models for structural biology. Before jumping into MLxBio, my background combined **fundamental research** experience with real-world **product impact**. I'm the first author of Scalable Inception Graph Network, now **used in production by AirBnB**, and Temporal Graph Network, which has received over **750 Github stars**. While at Twitter, I successfully developed multiple ML models for the "**who-to-follow**" recommendation feature, resulting in an increase of over **1.2M follow actions daily**.

Education

Imperial College London *Oct 2020 – Feb 2024*
Ph.D. in Computer Science – Supervisor: Prof. Michael Bronstein

- My research focused on solving the challenges preventing the success of **Graph Neural Networks** on real-world applications
- My work has been published at top conferences including ICML, ICLR, AAAI and RecSys

University of Cambridge *Oct 2018 – Jun 2019*
Master in Advanced Computer Science – Supervisor: Prof. Pietro Liò

- Graduated with **Distinction** (**4.0 GPA**)
- Thesis: "**Graph Deep Learning for ncRNA data**" (published at KDD 2019 GDL workshop)

Imperial College London *Oct 2015 – Jun 2018*
Bachelor in Computer Science

- Graduated with **First Class Honors** (**4.0 GPA**)

Experience

Machine Learning Researcher *NYC (remote)*
VantAI *Jan 2024 – Present*

- **Co-led development of Neo-1**, a state-of-the-art all-atom latent diffusion model for multimodal structure prediction and de-novo biomolecule generation.
  - **Model Design**: Led the design and exploration of model architectures, diffusion formulations, and inference methods, establishing efficient experimentation workflows.
  - **Large-scale Training**: Managed distributed training across 100+ GPUs, optimizing computational resources and model performance.
  - **Data Engineering**: Contributed to the development of high-quality, leakage-free biomolecular datasets (see Plinder, Pinder) used to train the model.
  - **Mentoring**: Mentored and supervised 3 PhD interns who made substantial contributions to the project.
  - **Cross-functional Collaboration**: Partnered closely with computational biologists and chemists to assess and enhance model applicability and accuracy.

Machine Learning Researcher *London*
Twitter *Jun 2019 – Feb 2023*

- Worked on **applied** and **fundamental research** around **graph neural networks**
- **Scaled** these technologies to graphs with **tens of billions of edges**
- Developed an embedding-based model to recommend users "who to follow", leading to an **increase of more than 1M follow actions** per day on the platform

Machine Learning Engineer *London*
Fabula AI *Mar 2019 – Jun 2019*

- Researched and developed **novel graph deep learning models** for **fake news classification**

- Fabula AI was acquired by Twitter in June 2019

Software Engineering Intern *California*
Google *Jul 2017 – Oct 2017*

- Worked in the **Google Play Store** infrastructure team

Selected Publications

- *E. Rossi et al.* Temporal Graph Networks for Deep Learning on Dynamic Graphs. *ICML 2020 GRL Workshop.* arXiv:2006.10637, blog. A new model for deep learning on dynamic graphs (¿300 citations, ¿700 Github stars). Adopted by memgraph (graph analytics company) and used by GraphCore to benchmark their hardware.
- *E. Rossi*, F. Frasca* et al.* SIGN: Scalable Inception Graph Neural Networks. *ICML 2020 GRL Workshop.* arXiv:2004.11198, blog. First graph deep learning model to scale to graphs with billions of edges (¿200 citations). Used in production by Airbnb.
- *B. Chamberlain, S. Shirobokov, E. Rossi et al.* Graph Neural Networks for Link Prediction with Subgraph Sketching. *ICLR 2022, Oral (top 5%).* arXiv:2209.15486. First GNN model to scale link prediction to graphs with millions of nodes thanks to sketching, a hashing technique which enables efficient computation of subgraph statistics.
- *E. Rossi et al.* On the Unreasonable Effectiveness of Feature propagation in Learning on Graphs with Missing Node Features. *Learning on Graphs Conference 2022.* arXiv:2111.12128, blog. First model to enable scalable graph learning with partially missing node features.

Projects and Awards

LeadTheFuture *Jun 2018 – Mar 2023*
Co-Founder

- **LeadTheFuture** is a **non-profit mentoring organization** and the **largest merit-based STEM community in Italy**, with more than **200 mentors** (engineers, researchers, and entrepreneurs) and **500 selected mentees**
- Our work has been featured in Forbes

Talks and Blog Posts

- I have written a series of blog posts on my research (full list) and have been invited to give over 10 talks (full list)